REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
29-MAY-2003	Conference Proceedings, (refereed)	

4. TITLE AND SUBTITLE
Spatio-Temporal Data Access For Information-Based Decision Making

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
ROY VICTOR LADNER   ELIZABETH G. WARNER   FREDERICK E PETRY

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER
74-6731-02

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Research Laboratory
Marine Geoscience Division
Stennis Space Center, MS 39529-5004

8. REPORTING ORGANIZATION REPORT NUMBER
NRL/PP/7440--03-1018

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217

10. SPONSOR/MONITOR'S ACRONYM(S)
ONR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release,distribution is unlimited

20031015 028

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Supporting large volumes of multi-dimensional data is an inherent characteristic of modem database applications, such as Geographical Information Systems (GIS), and Image and Multimedia Databases. Such databases need underlying systems with extended features like query languages, data models, and indexing methods, as compared to traditional databases, mainly because of the complexity of representing and retrieving data. In particular, spatio-temporal information systems must efficiently store and access potentially very large quantities of spatial and temporal data so we have developed index structures that can capture the time varying nature of moving objects, namely spatio-temporal structures. This paper describes the design of a query system for such data.

15. SUBJECT TERMS
multi-dimensional data, query languages, data models, and indexing methods

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON / 19b. TELEPHONE NUMBER *(Include area code)*
unclassified	unclassified	unclassified	Unlimited	5	Roy Ladner / 228-688-4679

Spatio-Temporal Data Access for Information-Based Decision Making

Roy Ladner, Elizabeth Warner, Fred Petry
Naval Research Laboratory
Stennis Space Center, MS, USA
{rladner, ewarner, fpetry}@nrlssc.navy.mil

Abstract - **Supporting large volumes of multi-dimensional data is an inherent characteristic of modern database applications, such as Geographical Information Systems (GIS), and Image and Multimedia Databases. Such databases need underlying systems with extended features like query languages, data models, and indexing methods, as compared to traditional databases, mainly because of the complexity of representing and retrieving data. In particular, spatio-temporal information systems must efficiently store and access potentially very large quantities of spatial and temporal data so we have developed index structures that can capture the time varying nature of moving objects, namely spatio-temporal structures. This paper describes the design of a query system for such data.**

I. Introduction

Supporting large volumes of multi-dimensional data is an inherent characteristic of modern database applications, such as Geographical Information Systems (GIS), and Image and Multimedia Databases. Such databases need underlying systems with extended features like query languages, data models, and indexing methods, as compared to traditional databases, mainly because of the complexity of representing and retrieving data. In particular, spatio-temporal information systems must efficiently store and access potentially very large quantities of spatial and temporal data so we have developed index structures that can capture the time varying nature of moving objects, namely spatio-temporal structures. This paper describes the design of a query system for such data.

The future requires military operations and intelligence communities to more heavily rely on Internet-based solutions for the delivery of MetOc data and products to the warfighter in an automated manner. These issues are being addressed by Tactical Environmental Data Services (TEDServices) [1]. TEDServices is being engineered by the Naval Research Laboratory, the Naval Oceanographic Office and the Naval Undersea Warfare Center, with sponsorship from Space and Naval Warfare Systems Command (SPAWAR) PMW-155. The Naval Research Laboratory's Geospatial Information DataBase System (GIDB) serves as the prototype for the system's components. The GIDB System has been under development and testing since 1994 under various sponsors with emphasis on Internet delivery of geospatial data from heterogeneous sources [2, 3].

II. Geospatial Information DataBase

The Digital Mapping, Charting and Geodesy Analysis Program (DMAP) at the Naval Research Laboratory has been actively involved in the development of digital geospatial mapping and analysis systems. This work started with the Geospatial Information Database (GIDB™), an object-oriented, CORBA-compliant spatial database capable of storing multiple data types from multiple sources. Data is accessible over the Internet via a Java Applet [4].

The GIDB includes an object-oriented data model, an object-oriented database management system (OODBMS) and various analysis tools. While the model provides the design of classes and hierarchies, the OODBMS provides an effective means of control and management of objects on disk such as locking, transaction control, etc. The OODBMS in use is Ozone, an open-source database management system [5]. This has been beneficial in several aspects. Among these, access to the source code allows customization and there are no costly commercial database licensing fees on deployment. Spatial and temporal analysis tools include query interaction, multimedia support and map symbology support. Users can query the database by area-of-interest, time-of-interest, distance and attribute. For example, statistics and data plots can be generated to reflect wave height for a given span of time at an ocean sensor. Interfaces are implemented to afford compatibility with Arc/Info, Oracle 8i, Matlab, and others.

The object-oriented approach has been beneficial in dealing with complex spatial data, and it has also permitted integration of a variety of raster and vector data products in a common database. Some of the raster data include satellite and motion imagery, Compressed ARC Digitized Raster Graphics (CADRG), Controlled Image Base (CIB), jpeg and video. Vector data includes Vector Product Format (VPF) products from the National Imagery and Mapping Agency (NIMA), Shape, real-time and in-situ sensor data and Digital Terrain Elevation Data (DTED). The VPF data includes such NIMA products as Digital Nautical Chart (DNC), Vector Map (VMAP), Urban Vector Map (UVMAP), Digital

Topographic Data Mission Specific Data Sets (DTOP MSDS), and Tactical Oceanographic Data (TOD).

Over the years, the system has been expanded to include a communications gateway that enables users to obtain data from a variety of data providers distributed over the Internet in addition to the GIDB. These providers include Fleet Numerical Meteorology and Oceanography Center (FNMOC), USGS, Digital Earth/NASA, and the Geography Network/ESRI. A significant FNMOC product is the Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS) data. The atmospheric components of COAMPS are used operationally by the U.S. Navy for short-term numerical weather prediction for various regions around the world. Our communications gateway provides a convenient means for users to obtain COAMPS data and incorporate it with other vector and raster data in map form. The gateway establishes a well-defined interface that brings together such heterogeneous data for a common geo-referenced presentation to the user. An illustration of the interface for a typical data request is shown in Figure 1.

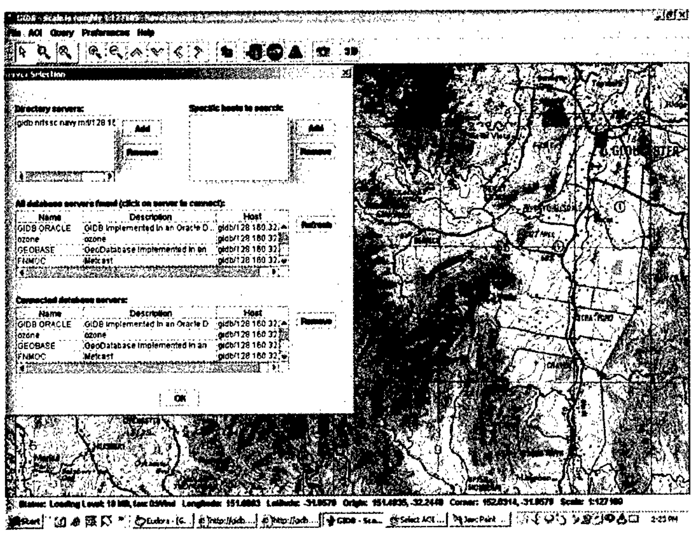

Figure 1: Example of GIDB Interface

III. TEDServices

Generally, TEDServices is a new, scaleable and modular environmental data repository, designed to support Warfighters, Weapon Systems, and Expert MetOc data Users. It includes a middleware infrastructure that enables the interoperable transport and transform of data, consistent with WGS84 datum and universal time coordinate, to and from a data repository, facilitated by a MetOc/Mission Rules Based Data Ordering scheme (MRBDO).

TEDServices provides new Web Services architecture within the Oceanographer of the Navy's (N096) Operational Concept 2002. TEDServices provides a Data Oriented Service (as defined by the Navy Enterprise Application Developers Guide – NEADG) that additionally supports the management and bi-directional transport of meteorological, oceanographic and environmental information. TEDServices offers a lightweight, forward deployed data cache, which offers warfighters, MetOc professionals, TDAs/applications and weapon systems immediate access to the Virtual Natural Environment (VNE), a 4-dimentional representation of the User define battle-space environment. TEDServices Clients will use a new MetOc/Mission Rules Based Data Order (MRBDO) process to subscribe to relevant data by mission, platform, TDA/application, parameter or product. The design tenants of TEDServices include: Data Transport (to reduce bi-directional bandwidth use), Data Management (to simplify data ordering and forwarded deployed data administration), Data Representation (implementation of a unified Geospatial Coordinate Process), and DoD Joint Interoperability (supporting standards defined by the Joint MetOc Interoperability Board).

IV. TEDServices Components

Some of the features and technologies of TEDServices are a pure Java implementation for platform independence, with planned support for of the Joint MetOc Interoperability Board (Navy, Air Force and Army) XML Interface Standard (Joint MetOc Broker Language – JMBL) and remote

administration. The conceptual components of TEDServices are shown in Figure 2, and include GateWays, Local DataBrokers, Local DataStores, and Interface support. These are explained below.

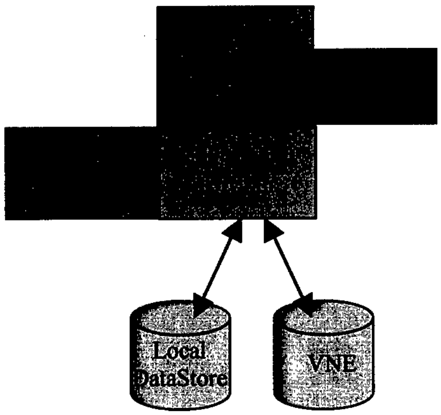

Figure 2: TEDServices Conceptual Components

The Local Data Broker (LDB) embodies the "smarts" of the system to pre-stage needed data at a particular location. It "knows" how to contact other TEDServices GateWays to request needed MetOc parameters over particular areas of interest. The LDB also monitors data usage and cancels further delivery of data not being used. The LDB also works to mitigate redundant reach-back requests for the same data by multiple users by caching data so that it is available for multiple users. The Data GateWay incorporates the software that streamlines the process of integrating data from heterogeneous sources to a common format and uniform datum, projection and universal time coordinate. A common format for all data within TEDServices simplifies format transformations to end-users. A MetOc/Mission Rules Based Data Ordering (MRBDO) system allows data requests to be aligned with relevant mission specific packages and platforms. This reduces the likelihood of requests for data that is not really essential to a particular task. It also offers a means of simplifying training. In TEDServices, GateWays communicate with other GateWays to forward deploy needed data to the end-user. User applications access data only from their local GateWay

V. Large Scale Data Transfer

Large scale data transfer can be difficult when network communications are unstable. TEDServices employs Resumable Object Streams (ROS) for all data traffic between major components of TEDServices across the network to achieve fail-safe data transportation under these conditions. ROS allows either the client or server side of a request to lose network connection, regain it, and the request will continue where it left off. In the event of a server shutdown and restart, server side processing of requests does not require the client to resend the request. Retransmission of the previously transmitted portion is not necessary in either case. Data requests can still be wrapped in compression and/or encryption.

ROS is not benchmarked to be any slower than non-ROS communication. Saving process state information to disk does not add significant overhead. The ROS transmission controls add almost no overhead to the communication (approximately 13 bytes). ROS works for broker to broker and broker to client applications.

In addition to ROS, a dynamic packet compression scheme (LPAC) was developed in conjunction with GMAI under a Small Business Innovation Research (SBIR) Phase II program, under the direction of the MetOc Systems Program Office at SPAWAR. LPAC provides higher lossless compression ratios than data compression methods currently favored by MetOc data users for large gridded data sets. Data is compressed prior to network transmission. It is also stored in the compressed format and uncompressed only on extraction to end-users. A Java-implementation DataBlade was transitioned to TEDServices by Barrodale Computing Services (BCS), under the direction of the MetOc Systems Program Office to provide the methods for complex extractions from these datasets.

VI. CASP – Sharing Applications' State

The Collaborative Application Sharing Process (CASP) is implemented in TEDServices to enable application users to share the state of their applications in a peer-to-peer manner. CASP will alleviate some of the planning requirements from the field and place them at a "center of expertise" where experts can perform some of the less time-critical planning and provide results to the field. CASP effectively allows various operational units to share information, without it having to be completely regenerated at a receiving site. This allows, in a U.S. Navy setting, heightened situational awareness in a distributed environment. CASP allows users to send to TEDServices a Java object that encapsulates the state of their application. This state is stored within TEDServices in a non-application specific manner that allows the objects to be disseminated to interested parties. When retrieved from TEDServices, the users may open the CASP object and restore the remote application's state on their own application, make any appropriate modifications and then submit the object back to TEDServices for further dissemination and sharing. Applications currently setup to use CASP include the Naval Integrated Tactical Environmental Subsystem II (NITES II) Object Oriented Redesign (OOR) and the Joint MetOc Viewer (3.6+).

VII. Environmental Decision Analysis

We are developing approaches for spatial data mining in an environment in which there is considerable concern about the development of ways for processing large amounts of spatio-temporal data especially of oceanographic and littoral regions and including meteorological information. Our plan is to integrate the data mining techniques into the GIDB. The ultimate goal is to provide knowledge-enhanced information to decision tools that will be used by US Navy and Marine planners [6].

Some of our approaches have included predictive modeling, attribute generalization and association rules for fuzzy spatial data and are more fully described in [7].

Attribute generalization, for example, is intended to provide a generalization or summarization of some potentially relevant aspects of the data being considered. This attribute-oriented induction approach produces a generalized representation by either attribute removal or attribute generalization. After this step the processed data is aggregated by merging identical tuples in the database and counting the number of tuples merged to indicate significance. Attributes are removed if there is no hierarchy for the attribute or if it can be expressed in terms of higher-level concepts of other attributes. Attribute generalization examines an attribute to ascertain if there are too large a number of distinct values (exceeding a given threshold). Then if a generalization hierarchy is available for this attribute, it is generalized and the common tuples merged.

We have applied this technique to sea bottom data from 10 locations (such as areas in the Philippines, Mediterranean, Persian Gulf, etc.). Here the intended application was to characterize various sea bottom areas for the planning of a mine deployment/hunting mission. The spatial data was queried to formulate the files from which the attribute generalization was done. The basic query was on bottom sediment classification as this was the major characteristic of interest to experts. The data consisted of polygons of the bottom types as classified. Depth was an estimate, and depth and area were binned into three categories.

We now consider particular aspects of TEDServices that address the needs of the warfighter with regard to visualization of battle space environmental impacts. Warfighters, such as special operations forces, require the ability to define a mission comprised of multiple components that are executed over event time and operational area. They also have the need to access a micro-forecast of the operation area that is shared throughout the planning process. Significant to mission success is the visualization of environmental impacts derived from the use of the micro-forecasts on the operational area fused with satellite images, dimensioned by event time. Finally, there is the presentation of the environmental impact mission plan to the warfighter in a physical format compatible with applications such as FalconView.

Environmental decision analysis requires environmental data – meteorological and oceanographic data containing relevant mission parameters. Then, decision analysis is performed in the form of red, green, yellow thresholds for parameters based on doctrinal guidelines. The GIDB provides the front-end thresholding and visualization engine connected to TEDServices and other sources available through the GIDB Portal. Data available through TEDServices includes data from the Naval Oceanographic Office, FNMOC and NPMOC-SD. Other sources available through the GIDB include NIMA data.

The integration process requires two stages. First, there is conversion to a uniform spatial resolution. Then, the data must be normalized to time. These enable an appropriate overlay for appropriate visual presentation to the user.

A mission may require several different platforms at any of the stages. The general approach to conduct environmental decision analysis involves use of doctrine-based rules for platforms at each of these stages. We can accept user input to define thresholds. Then we adapt to mission scenarios to determine changes in environment across time. That is,

1. Specified time and bounds – identify red, green and yellow areas;
2. Specified bounds and start time for mission – identify red, green, and yellow areas for each mission stage.

Finally, we visually present the results to the user in a time step or animated loop. We also allow the user to drill down to evaluate data values during the visual presentation.

A typical example of a query to support amphibious operations may involve parameters such as:

1. Wave Height
2. Wind
3. Water Clarity
4. Tides
5. Temperature

and many others. Some of the parameters are relevant for specific capabilities such in the case of parachute operations ceiling requirements for HAHO (High altitude – high opening) and HALO (High altitude – low opening), or surf conditions for combat swimmers. Parameter ranges have been operationally established and classified for many cases such those described in Table 1 for Ground Operations.

Table 1: Operational Parameter Ranges for Ground Operations

Ground Operation Parameter	Green – Favorable No Degradation	Amber – Marginal Degraded	Red – Unfavorable Severely Degraded
Visibility	>1 nm	.5 nm – 1 nm	< .5 nm
Temperature	20 F– 85 F	-15 F – 20 F or 85 F – 95 F	< -15 F and > 95 F
Lunar Illumination	< 75%	---	> 75%
Precipitation	Light	Moderate	Heavy
Ground Condition	Dry	Moist	Wet/Snow/Ice On Roads
Wind	< 25 knots	25 – 50 knots	> 50 knots
Wind Chill	> 20 F	-14 F – 20 F	≤-15 F

For a query the user selects the desired parameters and then assigns red, green and amber thresholds either from given specifications as in Table 1 or their own specialized knowledge. Three distinct spatio-temporal query types are then possible:

1) For a specified time T and bounding box the model will:

- Identify the best zone of operations
- Rank the located zones

2) For a specified time range T1 – T2, and a specified operational zone the model will:

- Identify the best time for the operation
- Rank the identified times

3) For a specified bounding box and no input search time the model will:

- Identify the best time and best zone for the operation
- Rank the times in order
- Rank the operation zones in order.

VIII. Summary

We have discussed issues pertinent to the Internet delivery and visualization of environmental data for network centric warfare. This presents many issues that must be addressed in order to effectively collect, share and access data and information from heterogeneous sources. Meteorological and oceanographic data presents many issues pertinent to access and retrieval of such data from heterogeneous sources in a distributed system. We addressed particular aspects of the delivery of environmental data that address the needs of special operations forces with regard to battle space environmental impacts.

IX. Acknowledgments

We would like to acknowledge the MetOc Systems Program Office, SPAWARSYSCOM PMW155, for sponsoring this research, and the TEDServices Engineering Working Group for their contributions to TEDServices.

X. References

[1] TEDServices: https://teds.navy.mil.

[2] Cobb, Maria, Foley, Harold, Wilson, Ruth, Chung, Miyi, Shaw, Kevin, "An OO Database Migrates to the Web," IEEE Software, pages 22-30, Volume 15, Number 3, May/June 98.

[3] D. Needham, R. Wilson, K. Shaw, "An Analysis of a CORBA-Based Approach to Accessing Geospatial Information via the Internet", 20th IEEE International Performance, Computing, and Communications Conference (IPCCC 2001), Phoenix, AZ, April 4-6, 2001, pp. 236-243.

[4] Chung, M., R. Wilson, R. Ladner, T. Lovitt, M. Cobb, M. Abdelguerfi, K. Shaw, "The Geospatial Information Distribution System (GIDS)", Chapter 17 in book, Succeeding with Object Databases, Dr. Akmal B. Chaudhri, editor, John Wiley & Sons, publishers, 2001.

[5] Ozone: http://www.ozone-db.org.

[6] Dr. Roy Ladner, F. Petry, M. Cobb, "Fuzzy Set Approaches to Spatial Data Mining of Association Rule", Transactions in GIS, Special Issue: Incorporating Fuzzy Sets in GIS, Vol 7, Issue 1, pp. 123-138, January 2003.

[7] Dr. Roy Ladner, Fredrick E. Petry, Assessment of Spatial Data Mining Tools for Integration into an Object-oriented GIS (GIDB), Dexa 2002, Aix-en-Provence, France, September 2002.